عالیه! اینکه بخوای سطح مدل هوش مصنوعی پروژه‌ات رو حرفه‌ای‌تر کنی یعنی داری می‌ری سمت یک \*\*Voice Assistant واقعی و باهوش\*\*. اجازه بده برات یک مسیر \*\*حرفه‌ای و کاربردی\*\* طراحی کنم برای ارتقای مدل، با تمرکز روی:

---

## 🎯 هدف: ارتقاء سطح مدل AI

ما می‌خوایم:

1. از مدل‌های NLP پیشرفته‌تر استفاده کنیم (مثل BERT).

2. پیش‌بینی دقیق‌تر داشته باشیم.

3. مدل رو آموزش بدیم و به فرمت `.tflite` برای اندروید تبدیل کنیم.

---

## ✅ مرحله ۱: استفاده از مدل BERT برای Intent Classification

### 📌 چرا BERT؟

\* مدل‌های پیش‌ساخته‌ی مثل BERT درک بسیار دقیق‌تری از زبان دارند.

\* برای تشخیص فرمان‌های متنی (مثل "چراغ را خاموش کن") بسیار مناسب هستند.

\* در فارسی هم مدل‌های آماده BERT وجود داره!

---

## 🧠 انتخاب مدل BERT فارسی

از مدل [\*\*HooshvareLab/bert-fa-base-uncased\*\*](https://huggingface.co/HooshvareLab/bert-fa-base-uncased) استفاده می‌کنیم که برای زبان فارسی آموزش دیده.

---

## 🛠️ مرحله ۲: آموزش مدل در پایتون

### کتابخانه‌ها:

```bash

pip install transformers datasets scikit-learn torch

```

### کد نمونه آموزش:

```python

from transformers import BertTokenizer, BertForSequenceClassification, Trainer, TrainingArguments

from datasets import load\_dataset

import torch

# دیتاست خودت رو آماده کن (text و label)

dataset = load\_dataset("csv", data\_files={"train": "train.csv", "test": "test.csv"}, delimiter=",")

tokenizer = BertTokenizer.from\_pretrained("HooshvareLab/bert-fa-base-uncased")

def tokenize\_function(example):

return tokenizer(example["text"], padding="max\_length", truncation=True)

tokenized\_datasets = dataset.map(tokenize\_function, batched=True)

tokenized\_datasets = tokenized\_datasets.rename\_column("label", "labels")

model = BertForSequenceClassification.from\_pretrained("HooshvareLab/bert-fa-base-uncased", num\_labels=6)

training\_args = TrainingArguments(

output\_dir="./results",

evaluation\_strategy="epoch",

per\_device\_train\_batch\_size=8,

per\_device\_eval\_batch\_size=8,

num\_train\_epochs=3,

save\_total\_limit=1,

logging\_dir='./logs',

)

trainer = Trainer(

model=model,

args=training\_args,

train\_dataset=tokenized\_datasets["train"],

eval\_dataset=tokenized\_datasets["test"],

)

trainer.train()

trainer.save\_model("trained-bert-fa")

tokenizer.save\_pretrained("trained-bert-fa")

```

---

## 🔄 مرحله ۳: تبدیل مدل به TFLite برای اندروید

BERT برای اندروید باید به `.tflite` تبدیل بشه. پیشنهاد من:

### روش ساده:

1. آموزش مدل با `onnx` یا `TF SavedModel`.

2. استفاده از `transformers` + `onnx` + `tf2onnx` یا `torch.onnx` → TFLite

### مسیر سریع:

```bash

pip install optimum

from optimum.exporters.onnx import main\_export

```

یا اگر با TensorFlow آموزش بدی، مستقیم با:

```bash

import tensorflow as tf

model = tf.keras.models.load\_model("saved\_model")

converter = tf.lite.TFLiteConverter.from\_saved\_model("saved\_model")

tflite\_model = converter.convert()

```

---

## 📲 مرحله ۴: استفاده از مدل در اندروید

بعد از تبدیل مدل به `.tflite`:

### Kotlin – بارگذاری مدل:

```kotlin

val tflite = Interpreter(loadModelFile("model.tflite"))

```

### آماده‌سازی ورودی:

توکنایزر باید همون باشه که تو آموزش استفاده شده:

\* توکنایزر رو از پایتون استخراج کن (`tokenizer.json`) و با خودت بیار تو اندروید.

\* ورودی رو به ترتیب `[CLS] input tokens [SEP]` و با padding طول ثابت آماده کن.

---

## ⚡ مرحله ۵: استفاده از صدای کاربر

\* تبدیل صوت به متن → ارسال متن به مدل → پیش‌بینی کلاس → ارسال فرمان به سیستم.

---

## 📁 ساختار حرفه‌ای نهایی

```

project/

│

├── backend/

│ ├── train\_model.py

│ ├── tokenizer/

│ ├── trained\_model/

│ └── export\_model.py

│

├── android/

│ ├── app/

│ │ ├── assets/

│ │ │ ├── model.tflite

│ │ │ └── tokenizer.json

│ │ └── TFLiteHelper.kt

│

└── README.md

```

---

## 💡 پیشنهادهای حرفه‌ای تکمیلی

| ویژگی | ابزار |

| --------------------- | ------------------------- |

| مدل فارسی با دقت بالا | BERT (Hooshvare) |

| توکنایزر هوشمند | BERT Tokenizer |

| شناسایی نیت و موجودیت | NLU + NER |

| اجرای سبک روی اندروید | TFLite + GPU Delegate |

| آموزش سفارشی | Datasets خودت با CSV/JSON |

---

### 📥 اگر دوست داشتی:

من می‌تونم:

\* کد آموزش کامل BERT رو برات بنویسم.

\* فایل آماده `.tflite` از مدل‌های فارسی برات درست کنم.

\* توکنایزر و JSONهاش رو خروجی بگیرم برای استفاده مستقیم در اندروید.

بگو کدوم مسیر رو انتخاب می‌کنی؟ یا اینکه فایل نمونه می‌خوای آماده برات بفرستم؟